

Analytic Investigations of Random Search Strategies for Parameter Optimization

Torsten Asselmeyer and Werner Ebeling
*Institut of Physics, Humboldt University Berlin,
 Invalidenstr. 110, D-10115 Berlin, Germany*

Several standard processes for searching minima of potential functions, such as thermodynamical strategies (simulated annealing) and biologically motivated selfreproduction strategies, are reduced to Schrödinger problems. The properties of the landscape are encoded in the spectrum of the Hamiltonian. We investigate this relation between landscape and spectrum by means of topological methods which lead to a possible classification of landscapes in the framework of the operator theory. The influence of the dimension d of the search space is discussed.

The connection between thermodynamical strategies and biologically motivated selfreproduction strategies is analyzed and interpreted in the above context. Mixing of both strategies is introduced as a new powerful tool of optimization.

PACS numbers: 05.40.+j, 05.90.+m, 03.65.Db

submitted to Phys.Rev. E

I. INTRODUCTION

The optimization problem appears in several fields of physics and mathematics. It is known from mathematics that every local minimum of a convex function defined over a convex set is also the global minimum of the function. But the main problem is to find this optimum. From the physical point of view every dynamical process can be considered in terms of finding the optimum of the action functional. The best example is the trajectory of the free point mass in mechanics which follows the shortest way between two points.

Let us assume one has successfully set up a mathematical model for the optimization problem under consideration in the form

$$U(x_1, x_2, \dots, x_d) \longrightarrow \text{Min} \quad (1)$$

where U is a scalar potential function (fitness function) defined over a d -dimensional vector space $X = \mathcal{L}\{x_1, \dots, x_d\}$. Let $x^{(0)}$ be the absolute Minimum of U which is the search target. Problems of this type are called parameter optimization. Typically the search space is high-dimensional ($d \gg 1$).

The idea of evolution is the consideration of an ensemble of searchers which move through the search space. As a illustrative example we consider the relation between the equation of motion in mechanics, obtained by variation of the action functional S to get a trajectory of minimal action, and the introduction of the probability distribution for all possible trajectories by a functional integral. Because of the weight factor $\exp(-S/\hbar)$ for every trajectory, the trajectory of minimal action has the highest probability. The equation of the probability distribution deduced from the functional integral is the diffusion equation for a free classical particle. The same idea is behind the attempt to describe optimization processes with the help of dynamical processes.

We will be concerned with the time evolution of an ensemble of searchers defined by a density $P(\vec{x}, t)$. The search process defines a dynamics

$$P(\vec{x}, t + \Delta t) = \mathbf{T}[P(\vec{x}, t); \Delta t] \quad (2)$$

with continuous or discrete time steps. An optimization dynamics \mathbf{T} is considered as successful, if any (or nearly any) initial density $P(\vec{x}, 0)$ converges to a target density $\lim_{\tau \rightarrow \infty} P(\vec{x}, \tau)$ which is concentrated around the minimum $x^{(0)}$ of $U(\vec{x})$. We restrict ourselves here to the case where \mathbf{T} is given as a second order partial differential equation. Among the most successful strategies are the class of thermodynamic oriented [1–3] and the class of biological oriented strategies [4–6]). Our aim is to compare on the basis of PDE-models thermodynamical and biologically motivated strategies by reducing both to equivalent eigenvalue problems. Further we introduce a model for mixed strategies and investigate their prospective power [7–9].

II. THERMODYNAMICAL STRATEGIES

At first we want to investigate the simplest case of an evolutionary dynamics known in the literature as “simulated annealing”. The analogy between equilibrium statistical mechanics and the Metropolis algorithm [10] was first discussed by Kirkpatrick et al. in [11]. There, an ensemble of searchers determined by the distribution $P(\vec{x}, t)$ move through the search space. In the following we consider only the case of a fixed temperature. Then the dynamics is given by the Fokker-Planck equation

$$\frac{\partial}{\partial t} P(\vec{x}, t) = \nabla D \cdot [\nabla P + \beta \nabla U P] = D \Delta P + D \nabla (\beta P \nabla U) \quad (3)$$

with the “diffusion” constant D , the reciprocal temperature β and the state vector \vec{x} . The stationary distribution $P_0 \sim \exp(-\beta U(\vec{x}))$ is equal to the extremum of functional (11) [12].

For the case $\beta = 0$ the complete analytical solution is known and may be expressed by the well-known heat kernel or Green’s function, respectively. The corresponding dynamics is a simulated annealing at an infinite temperature which describes the diffusion process. In this case the optimum of the potential will be found by a random walk, because the diffusion is not sensitive to the potential. The average time which the process requires to move from the initial state $\vec{x}(0)$ to the optimum $\vec{x}^{(0)}$ is given by

$$t_0 = \frac{1}{D} < (\vec{x}(0) - \vec{x}^{(0)})^2 > \quad (4)$$

We note that in several cases a generalization from the number $D = \text{const.}$ to the case where D is a symmetrical matrix is possible [12]. Further solvable cases can be extracted from the ansatz

$$P(\vec{x}, t) = \exp \left[-\frac{\beta U(\vec{x})}{2} \right] y(\vec{x}, t) \quad (5)$$

which after separation of the time and space variables leads to the eigenvalue equation

$$(\epsilon - H)\psi(\vec{x}) = D\Delta\psi(\vec{x}) + (\epsilon - V(\vec{x}))\psi(\vec{x}) = 0 \quad (6)$$

$$y(\vec{x}, t) = \exp(-\epsilon t)\psi(\vec{x})$$

with eigenvalue ϵ and potential

$$V(\vec{x}) = \frac{\beta^2}{4} D \nabla U \cdot \nabla U - \frac{\beta}{2} D \Delta U \quad (7)$$

This equation is the well-known stationary Schrödinger equation from quantum mechanics. Under the consideration of a discrete spectrum this leads to the general solution

$$P(\vec{x}, t) = \exp \left[-\frac{\beta}{2} U(\vec{x}) \right] \sum_{i=0}^{\infty} c_i \psi_i(\vec{x}) \exp(-\epsilon_i t) \quad (8)$$

To discuss more properties of equation (6), one has to introduce the concept of the Liapunov functional [13] defined by the formula:

$$\frac{d}{dt} y(\vec{x}, t) = -\frac{\delta}{\delta y} \mathcal{L}(y, \nabla y) \quad (9)$$

In the case of equation

$$\frac{d}{dt} y(\vec{x}, t) = D\Delta y(\vec{x}, t) - V(\vec{x})y(\vec{x}, t) \quad (10)$$

we obtain

$$\mathcal{L}(y, \nabla y) = \int_X L(y, \nabla y) d\text{vol}(X) \quad (11)$$

$$L(y, \nabla y) = \frac{D}{2} (\nabla y)^2 + \frac{1}{2} V y^2 \quad .$$

For the original equation (3), the construction of such functional is impossible.

We remark that the main difference between Schrödinger equation and thermodynamical strategies is given by the time-dependent factor in the solution. In quantum mechanics this set of factors $\exp(i\epsilon t)$ forms a complete basis in the Hilbert space of functions over t but in the solution of (3) this is not the case.

Because of the existence of an equilibrium distribution the first eigenvalue ϵ_0 vanishes and the solution is given by

$$P(\vec{x}, t) = \text{const.} \exp[-\beta U(\vec{x})] \quad (12)$$

That means that the equilibrium distribution is located around the optimum since the exponential $\exp(-U(x))$ is a monotonous function and the optimum is unchanged. In the limit $t \rightarrow \infty$ the distribution $P(\vec{x}, t)$ converges to the equilibrium distribution and the strategy successfully terminates at the optimum. But this convergence is dependent on the positiveness of the operator H defined in (6). Usually the Laplace operator is strictly negative definite with respect to a scalar product in the Hilbert space of the square integrable functions (L^2 -space). Therefore the potential alone determines the definiteness of the operator. We thus obtain

$$0 < \int_X y(x) V(x) y(x) d\text{vol}(X) \quad .$$

A sufficient condition for that is

$$\frac{\beta}{2} (\nabla U)^2 > \Delta U \quad (13)$$

which means, that the curvature of the landscape represented by ΔU must be smaller than the square of the gradient. Depending on the potential it thus is possible to fix a subset of X on which the operator H is positive definite.

Now we approximate the fitness function $U(\vec{x})$ by a Taylor expansion around the optimum including the second order. Because the first derivative vanishes one obtains the expression

$$U(x) = U_{\min} + \frac{1}{2} \sum_{i=1}^d a_i (\vec{x}_i - \vec{x}_i^{(0)})^2 \quad (14)$$

For $a_i > 0$ we get the simple harmonic oscillator which is solved by separation of variables. The eigenfunctions are products of Hermitian polynomials with respect to the dimension of the search space. Apart from a constant the same result is obtained in the case $a_i < 0$. A collection of formulas can be found in appendix A.

The approximation of the general solution (8) for large times leads to

$$P(\vec{x}, t) = c_0 \exp(-\beta U(\vec{x})) + c_1 e^{-\epsilon_1 t} \exp\left[-\frac{\beta}{2} U(\vec{x})\right] \psi_1(\vec{x}) + \dots \quad (15)$$

Because of the condition $\epsilon_2 > \epsilon_1$, the time $\tau = 1/\epsilon_1$ can be interpreted as relaxation time, i.e. the time for the last jump to the optimum. Even more interesting than the consideration of the time is the calculation of velocities. One can define two possible velocities. A first velocity $v^{(1)}$ on the fitness landscape and a second one $v_k^{(2)}$ in the k -th direction of the search space. The measure of the velocities is given by the time-like change of the expectation values of the vector x_k or the potential $U(\vec{x})$, respectively. With respect to equation (3) we obtain

$$v^{(1)} = -\frac{d}{dt} \langle U \rangle = D\beta \langle \nabla U \cdot \nabla U \rangle - D \langle \Delta U \rangle \quad (16)$$

and

$$v_k^{(2)} = -\frac{d}{dt} \langle x_k \rangle = D\beta \langle \nabla x_k \cdot \nabla U \rangle \quad (17)$$

The velocity $v^{(1)}$ depends on the curvature and the gradient. So we can deduce a sufficient condition for a positive velocity

$$\beta(\nabla U)^2 > \Delta U \quad (18)$$

which is up to a factor the same as condition (13). For the quadratic potential (14) one obtains

$$|x_i - x_i^{(0)}| > \frac{1}{\sqrt{a_i \beta}} \quad \forall i \quad (19)$$

This is a restriction to a subset of X .

In this case it is also possible to explicitly calculate the velocities for $a_i > 0$

$$v^{(1)} = \frac{2c_2}{\beta c_1} \epsilon_2 \exp(-\epsilon_2 t) \quad (20)$$

$$v_k^{(2)} = \sqrt{\frac{2}{\beta a_k}} \frac{c_1}{c_0} \epsilon_1 \exp(-\epsilon_1 t) \quad (21)$$

It is interesting to note that only the first two eigenvalues are important for the velocities and that both velocities vanish in the limit $t \rightarrow \infty$. Besides, the first velocity $v^{(1)}$ is independent of the parameter a_i except for special initial conditions where the factor c_k depends on the parameter a_i . The other case $a_i < 0$ is similar and can be found in appendix A.

III. FISHER-EIGEN STRATEGIES

In principle, the biologically motivated strategy is different from the thermodynamical strategy. Whereas in

the thermodynamical strategy the population size remains constant, it is changed with respect to the fundamental processes reproduction and selection in the case of biologically motivated strategies, but is kept unchanged on average. The simplest model with a similar behaviour is the Fisher-Eigen equation given by

$$\frac{\partial}{\partial t} P(\vec{x}, t) = [\langle U \rangle - U(\vec{x})] P(\vec{x}, t) + D \Delta P(\vec{x}, t) \quad (22)$$

$$\langle U \rangle(t) = \frac{\int U(x) P(\vec{x}, t) dx}{\int P(\vec{x}, t) dx}$$

In this case one can also form a Liapunov functional which satisfies the equation (9) similar to the thermodynamical strategy. One obtains the positive functional

$$\mathcal{L} = \int_X \left(\frac{D}{2} (\nabla P)^2 - \frac{1}{2} (\langle U \rangle - U) P^2 \right) d\text{vol}(X) \quad (23)$$

which also has the stationary distribution as an extremum.

By using the ansatz

$$P(\vec{x}, t) = \exp \left[\int_0^t \langle U \rangle(t') dt' \right] y(\vec{x}, t) \quad (24)$$

and the separating time and space variables, the dynamics reduces to the stationary Schrödinger equation

$$(\epsilon_i - H) \psi_i(\vec{x}) = D \Delta \psi_i(\vec{x}) + [\epsilon_i - U(\vec{x})] \psi_i(\vec{x}) = 0 \quad (25)$$

where ϵ_i are the eigenvalues and $\psi_i(\vec{x})$ are the eigenfunctions. This leads to the complete solution

$$y(\vec{x}, t) = \sum_i a_i e^{-\epsilon_i t} \psi_i(\vec{x}) \quad (26)$$

The difference to the thermodynamical strategy is given by the fact that the eigenvalue ϵ_0 in the case of the Fisher-Eigen strategy is a non zero value, i.e. the relaxation time is modified and one obtains

$$t_0 = \frac{1}{\epsilon_1 - \epsilon_0} \quad (27)$$

For the harmonic potential (14) the problem is exactly solvable for any dimension d and the solution is very similar to the thermodynamical strategy for $a_i > 0$. In the other case $a_i < 0$ we obtain a different problem known from scattering theory. If the search space is unbound the spectrum of the operator H is continuous. From the physical point of view we are interested in positive values of the potential or fitness function (14), respectively. This leads to a compact search space given by the interval $[-b, b]$ in each direction with $b = \sqrt{2U_{\min}/|a_i|}$. We now have to introduce boundary conditions. The most natural choice is to let the solution vanish on the boundary. As a result an additional restriction appears and the spectrum of the operator H is now discrete. A collection

of formulas connected to both eigenvalue problems can be found in appendix B.

The next step is the calculation of the velocities defined in the previous section. With respect to the Fisher-Eigen equation (22) one obtains

$$v^{(1)} = \langle U^2 \rangle - \langle U \rangle^2 - D \langle \Delta U \rangle \quad (28)$$

$$v_k^{(2)} = \langle x_k U \rangle - \langle x_k \rangle \langle U \rangle \quad (29)$$

For the case of a quadratic potential all velocities can be calculated from the solution and with $a_i > 0$ the following expansion is possible:

$$v^{(1)} = 4D\epsilon_0 + O(e^{-(\epsilon_2 - \epsilon_0)t}) \quad (30)$$

$$v_i^{(2)} = \sqrt{\frac{18D}{a_i}} \frac{c_1}{c_0} \epsilon_1 e^{-(\epsilon_1 - \epsilon_0)t} + O(e^{-(\epsilon_3 - \epsilon_0)t}) \quad (31)$$

These velocities are very similar to the velocities of the thermodynamical strategy. So it is very difficult to decide whether or not the thermodynamical strategy is faster than the biologically motivated one or vice versa. This decision depends on the particular circumstances. The thermodynamical strategy is faster than the biologically motivated one in a landscape with a slight curvature and widely extended hills whereas the biological strategy needs a landscape with a high curvature and more localized hills to be faster than the thermodynamical strategy.

We note the interesting fact, that there exist special tunnel effects connected with minima of equal depth and shape. The corresponding spectrum of the Hamiltonian shows degenerated eigenvalues. Then under the condition of overlapping distributions located in different minima an tunnelling with high probability between these minima is possible (see Fig. 1, dotted line). For the corresponding Boltzmann strategy the transformed potential does not admit this tunnelling effect (Fig. 1, solid line).

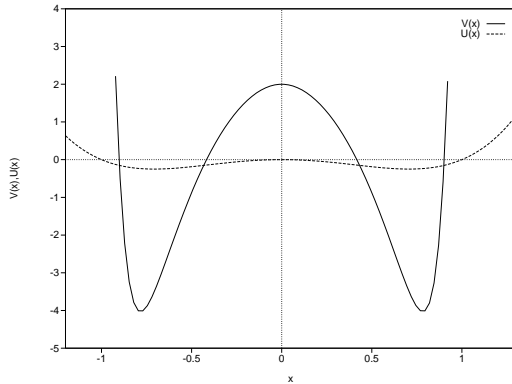


FIG. 1. The double well problem for Fisher-Eigen and Boltzmann

IV. MIXED BOLTZMANN–DARWIN STRATEGIES

The dynamic equations defining Boltzmann-type search and Fisher-Eigen type search contain a common term $D\Delta P$. Since both types of strategies have definite advantages and disadvantages it seems desirable to mix them. We defined the dynamics of a mixed strategy by

$$\frac{\partial}{\partial t} P(\vec{x}, t) = D\Delta P(\vec{x}, t) + \beta D \nabla (P \nabla U) \quad (32)$$

$$+ \gamma [\langle U \rangle - U(\vec{x})] P(\vec{x}, t) \quad (33)$$

For $\gamma = 0$ this dynamics reduces to a pure Boltzmann strategy and for $\beta = 0$ we obtain a Fisher-Eigen strategy. The mixed case may be treated by means of the ansatz

$$P(\vec{x}, t) = \exp \left[\gamma \int_0^t \langle U \rangle dt' - \frac{1}{2} \beta U(\vec{x}) \right] y(\vec{x}, t) \quad (34)$$

which leads to the explicit solution in the eigenfunctions of the problem

$$0 = D\Delta \psi_i(\vec{x}) + [\epsilon_i - E(\vec{x})] \psi_i(\vec{x}) \quad (35)$$

$$E(\vec{x}) = \gamma U - \frac{D\beta}{2} \Delta U + \frac{D\beta^2}{4} (\nabla U) \cdot (\nabla U) \quad .$$

For $\beta = 0$ and $\gamma = 1$ we end up with the case of Fisher-Eigen strategies and for $\gamma = 0$, $\beta > 0$ the Boltzmann dynamics is obtained. In this respect the mixed strategy is indeed a generalization of both cases (see Fig. 2).

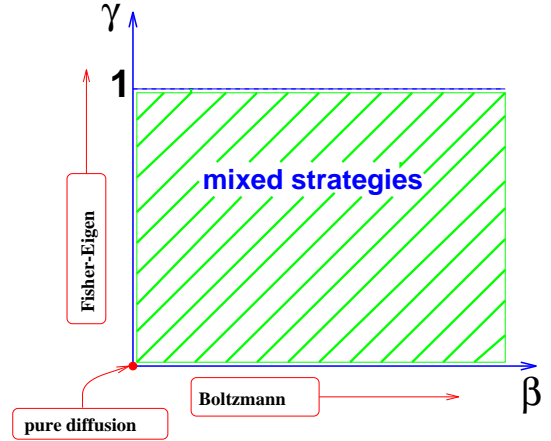


FIG. 2. The parameter dependence of the mixed strategy

The linearity of the differential equation leads to simple relations between the solutions and velocities. For instance the velocities for the thermodynamical and the biological strategy can be added with respect to the constant γ to get the velocities of the mixed strategy. For the solution of the problem one simply takes the solution of the thermodynamical strategy (A1) and redefines the coordinate ξ_k by

$$\xi_k = x_k \sqrt[4]{\frac{\beta^2 a_k^2}{4} + \frac{\gamma a_k}{2D}}.$$

So the mixed strategy combines the advantages of both strategies if one uses a criterion to choose the coefficients γ and β . The dependence on the 3 optimization parameters D , β and γ is a rather simple one. For a given D , the best strategy is evidently an increase of the β -value, since the γ -dependence is weak. However, if the “friction” βD is fixed by the conditions, the maximization of γ/β leads to the smallest relaxation time. In other words, if the loss βD is fixed, maximal γ (i.e. most competition) and minimal β (i.e. large temperature) is the best choice for a fast search of the minimum. In general the optimal search requires $\beta > 0$ and $\gamma > 0$. With other words, adding some amount of the “complementary” strategy is in most cases to be recommended. This was already found empirically in a earlier work [7,8].

V. GLOBAL PROPERTIES OF STRATEGIES

The study of the problem given by the potential (14) is the same as the study of the local properties of the landscape. It is known from the Lemma of Morse, that there is a coordinate system so that every function with non-degenerate critical points in the neighborhood U_p of one critical point p can be expressed as

$$f = f(p) - (y_1)^2 - \dots - (y_\lambda)^2 + (y_{\lambda+1})^2 + \dots + (y_d)^2$$

with (y_1, \dots, y_d) as coordinate system. The number λ is the index of the critical point. Now we use the compact space $X \cup \{\infty\}$ rather than $X = \mathbb{R}^d$ which in contrast to \mathbb{R}^d is homeomorphic to S^d , the d -dimensional sphere. Now every function over S^d is non-degenerate and we obtain for the number C_λ of critical points with index λ the Morse inequalities [14]

$$1 \leq C_0 \quad 1 \leq C_d \quad 0 \leq C_\iota \quad (36)$$

$$1 + (-1)^d = \sum_\lambda (-1)^\lambda C_\lambda \quad (37)$$

with $\iota = 1, \dots, d-1$. So we find that in the case of many maxima and minima the number of saddle points increases [15]. This is a global statement which only depends on the topology of the sphere S^d .

In the case of a compact, simple-connected subset of \mathbb{R}^d one obtains a similar result because the boundary of this subset is homeomorphic to a $d-1$ dimensional sphere and the Morse inequalities can be used again. If this subset has no boundary then we obtain

$$1 = \sum_\lambda (-1)^\lambda C_\lambda \quad (38)$$

and the result of [15] can be deduced again. The treatment of constraints in the optimization problem is connected with the occurrence of holes in the space. This

leads to a correction of the formulas (37) or (38), respectively. The number of saddle points increases with respect to the number of holes.

Next we want to study the influence of the landscape on the solutions in terms of the dynamics, which in our case is restricted to the influence of the potential or fitness function, respectively. The general form of this dynamics is given by

$$\frac{\partial}{\partial t} P(x, t) = -H P(x, t) \quad (39)$$

with the selfadjoint Operator

$$H_B = -D \nabla \cdot (\nabla + \beta \nabla U) \quad \text{Boltzmann} \quad (40)$$

$$H_{FE} = -\langle U \rangle + U - D \Delta \quad \text{Fisher-Eigen} \quad (41)$$

This operator equation has the formal solution

$$P(x, t) = \exp(-tH) P(x, 0) \quad (42)$$

acting on the initial condition. With the help of (24), the operator for the Fisher-Eigen strategy can be reduced to $H = U - D \Delta$. Then both operators are in the same class known as generalized Laplacian. The representation of this class from a unified point of view is possible. To this end we introduce the Dirac operator as a first order differential operator and regard the generalized Laplacian H as the square of the Dirac operator. It was shown in [16] that every generalized Laplacian can be represented by the square of a Dirac operator with respect to a suitable Clifford multiplication. For the two cases one obtains

$$D_B = i\gamma^\mu (\partial_\mu + \partial_\mu U) \quad \text{Boltzmann} \quad (43)$$

$$D_{FE} = \gamma^\mu (\partial_\mu + iA_\mu(x)) \quad \text{Fisher-Eigen} \quad (44)$$

where γ^μ are the Dirac matrices ($\mu = 1, \dots, d$) and $A_\mu(x)$ is a locally defined vector field with $[\gamma^\mu, \gamma^\nu] \partial_\mu A_\nu(x) = -iU(x)$. Together with the expression

$$D = \begin{pmatrix} 0 & \not\partial \\ \not\partial^* & 0 \end{pmatrix} \quad (45)$$

for every Dirac operator D , one can simple calculate the squares of the Dirac operators to establish

$$H_B = D_B^2 = \begin{pmatrix} \not\partial_B \not\partial_B^* & 0 \\ 0 & \not\partial_B^* \not\partial_B \end{pmatrix} \quad H_{FE} = D_{FE}^2 \quad (46)$$

with the adjoint operator $\not\partial^*$. This means that both problems can be described by the motion of a fermion in a field ∇U or $iA(x)$, respectively. The equilibrium state (or stationary state) is given by the kernel of the operator H which is the direct sum $\ker H = \ker \not\partial \oplus \ker \not\partial^*$. If the dimension d is even, which is always true for the high-dimensional case, this splitting can be introduced by the product of all γ -matrices usually denoted by γ^5 . Because of the compactness of the underlying space S^d , the

kernel of H is finite-dimensional and the spectrum is discrete. We note, that the spectrum of both, $\partial\partial^*$ and $\partial^*\partial$, is equal up to the kernel. So the interesting information about the problem is located in the asymmetric splitting of the kernel $\ker H$. The physical interpretation is given by an asymmetric ground state of the problem which is only connected to the geometry of the landscape. In more mathematical terms both Dirac operators are described as covariant derivatives with a suitable connection. Together with the fiber bundle theory [17] and the classification given by the K-theory [18] one obtains a possible classification of the fitness landscapes in dependence of convergence velocities given by the periodicity of the real K-theory with period 8. Each of these 8 classes describes a splitting of the kernel $\ker H$ and leads to a different velocity. A complete description of this problem will be published later on.

VI. CONCLUSIONS

In physics the most classical dynamical processes follow the principle of minimization of a physical quantity which often leads to an extremum of the action functional. This problem frequently has a finite number of

solutions given by the solutions of a differential equation also known as the equation of motion. Investigating this fact in relation to optimization processes, one obtains in the simplest case the thermodynamical and the biological strategy. The description is given by the distribution of the searcher and a dynamics of the distribution converging to an equilibrium distribution located around the optimum of the optimization problem. With the help of the kinetics and the eigenfunction expansion, we investigated both strategies in view of the convergence velocity. In principle both strategies are equal because one obtains a stationary Schrödinger equation. But, the main difference is the transformation of the fitness function (or potential) from $U(x)$ to $V(x)$ (see (7)) in the case of the thermodynamical strategy. The difference of both strategies leads to the idea of adding a small amount of the “complementary” strategy in order to hope for an improvement. The difference in the velocity on the one hand and the similarity in the equation on the other look like a unified treatment of both strategies under consideration. This is represented in the last section in the formalism of fiber bundles and heat kernels to get the interesting result, that up to local coordinate transformations the strategies are split into 8 different classes.

APPENDIX: A - THERMODYNAMICAL STRATEGY

For the case of quadratic potentials (14) the problem (3) may be solved explicitly. We get the eigenvalues

$$\epsilon_{n_1 \dots n_d} = U_{min} + \sum_{i=1}^d a_i \beta D \left(n_i + \frac{1}{2} \right) \quad n_i = 0, 1, 2, \dots$$

and the eigenfunctions

$$\begin{aligned} \psi_{n_1 \dots n_d}(x_1, \dots, x_d) &= \prod_{i=1}^d \psi_{n_i}(x_i) \\ \psi_{n_i}(x_i) &= \exp\left(-\frac{\beta a_i}{4} x_i^2\right) H_{n_i}\left(\sqrt{\frac{\beta a_i}{2}} x_i\right) \end{aligned}$$

which lead to the solution

$$P(\vec{x}, t) = \sum_{i=n_1+\dots+n_d} c_i \prod_{k=1}^d [\exp(-\xi_k^2) H_{n_k}(\xi_k)] \exp(-\epsilon_i t) \quad (A1)$$

with $\xi_k = x_k \sqrt{\beta a_k/2}$. Next we have to fix initial conditions for this problem. At first one starts with a strong localized function, i.e. with a delta distribution.

$$P(\vec{x}, 0) = \prod_{i=1}^d \delta(x_i - x_i^{(0)})$$

Because of the relation:

$$\int \exp(-\xi^2) H_n(\xi) H_m(\xi) d\xi = \delta_{mn} \int \exp(-\xi^2) H_n^2(\xi) d\xi = \delta_{mn} N_n = \delta_{mn} 2^n (n!) \sqrt{\pi} \quad (A2)$$

we obtain for the coefficients

$$c_k = \prod_{i=1}^d \frac{2}{N_{n_i} \beta a_i} H_{n_i} \left(\sqrt{\beta a_i} 2x_i^{(0)} \right) \quad . \quad (\text{A3})$$

with $k = n_1 + n_2 + \dots + n_d$. In the case of the full symmetry $a = a_1 = a_2 = \dots = a_d$ we can calculate the radial problem to obtain the eigenvalues

$$\epsilon = a\beta Dd \left(k + \frac{1}{2} \right)$$

with d as the dimension of the landscape. The calculation of the velocities leads to two cases for the potential (14):

1. $a_i > 0$

$$v^{(1)} = \frac{2c_2}{\beta c_1} \epsilon_2 \exp(-\epsilon_2 t) \quad (\text{A4})$$

$$v_k^{(2)} = \sqrt{\frac{2}{\beta a_k}} \frac{c_1}{c_0} \epsilon_1 \exp(-\epsilon_1 t) \quad (\text{A5})$$

2. $a_i < 0$

$$v^{(1)} = \frac{2}{\beta N_1} \sum_{k=0}^{\infty} c_{2k} \exp(-\epsilon_{2k} t) \epsilon_{2k} \sum_{i=1}^d \frac{H_{2k+3}(b_i)}{(2k+1)(2k+2)(2k+3)} \prod_{\substack{j=1 \\ k_1+\dots+k_d=k}}^d \frac{H_{2k_i+1}(b_j)}{(2k+1)} \quad (\text{A6})$$

$$v_i^{(2)} = \sqrt{\frac{2}{\beta a_k}} \frac{1}{N_1} \sum_{k=0}^{\infty} c_{2k+1} \exp(-\epsilon_{2k+1} t) \epsilon_{2k+1} \frac{H_{2k+3}(b_i)}{(2k+2)(2k+3)} \prod_{\substack{j=1 \\ k_1+\dots+k_d=k}}^d \frac{H_{2k_i+1}(b_j)}{(2k+1)} \quad (\text{A7})$$

where

$$N_1 = \sum_{k=0}^{\infty} c_{2k} \exp(-\epsilon_{2k} t) \prod_{\substack{i=1 \\ k_1+\dots+k_d=k}}^d \frac{H_{2k_i+1}(b_i)}{(2k+1)}$$

is the normalization factor and $b_i = \sqrt{2U_m/a_i}$ is the interval length.

APPENDIX: B - BIOLOGICAL STRATEGY

For a harmonic potential (14) the problem (22) is exactly solvable for any dimension d . We get the eigenvalues

$$\epsilon_{n_1 \dots n_d} = U_{min} + \sum_{i=1}^d \sqrt{2a_i D} \left(n_i + \frac{1}{2} \right) \quad n_i = 0, 1, 2, \dots$$

and the eigenfunctions

$$\begin{aligned} \psi_{n_1 \dots n_d}(x_1, \dots, x_d) &= \prod_{i=1}^d \psi_{n_i}(x_i) \\ \psi_{n_i}(x_i) &= \exp \left(-\sqrt{\frac{a_i}{4D}} x_i^2 \right) H_{n_i} \left(\sqrt{\frac{a_i}{2D}} x_i \right) \end{aligned}$$

which lead to the solution

$$P(\vec{x}, t) = \sum_{i=n_1+\dots+n_d} c_i \prod_{k=1}^d \left[\exp \left(-\frac{\xi_k^2}{2} \right) H_{n_k}(\xi_k) \right] \exp(-\epsilon_i t) \quad (\text{B1})$$

with $\xi_k = x_k \sqrt[4]{a_k/(2D)}$.

We now come to the problem of the maximum, i.e. the potential (14) with $a_i < 0$. The solution for one dimension (direction i) is simply obtained as

$$\psi_{\lambda_i}(x_i) = D_{-iB_i-\frac{1}{2}} \left(\sqrt[4]{\frac{a_i}{2D}} e^{-i\pi/4} x_i \right) \quad (B2)$$

with $B_i = \lambda_i/(2\sqrt{Da_i})$ and $D_k(x)$ as parabolic Bessel functions depending on the parameter B_i . In practice one is interested in positive values of the fitness function or potential which leads to a restriction of the search space X to the a hypercube with length $b_i = \sqrt{2U_m/a_i}$ in every dimension. We claim that the solution vanishes at the boundary of the hypercube. This restriction leads to a discrete spectrum. The zeros of the parabolic Bessel function are given by the solution of the equation

$$\frac{1}{4} \left(\rho_{ki} \sqrt{\rho_{ki}^2 - 1} - \text{arccosh } \rho_{ki} \right) = \frac{(-A_n)^{3/2}}{6\sqrt{\frac{U_m+n_i}{2\sqrt{a_i}D}}} \quad (B3)$$

where the coefficients A_i can be found in the book [19] and the zeros are defined by $c = 2\sqrt{a_i}\rho$. For the eigenvalues ϵ_k in (25) one obtains

$$\epsilon_n = \sum_{\substack{i=1 \\ n_1+\dots+n_d=n}}^d n_i \quad n_i = \frac{U_m}{\rho_{ki}^2} \sqrt{\frac{D}{a_i}} - U_m \quad . \quad (B4)$$

The calculation of the velocities is very lengthy and one obtains

1. $a_i > 0$ (minimum):

$$v^{(1)} = \frac{1}{N_1} \sum_{i=1}^d \sqrt{\frac{Da_i}{2}} \left(c_0 \epsilon_0 e^{-\epsilon_0 t} + \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(2k)!}{k!} (2k^2 + k + \frac{3}{2}) c_{2k} \epsilon_{2k} e^{-\epsilon_{2k} t} \right) \quad (B5)$$

$$v_i^{(2)} = \sqrt{\frac{2D}{a_i}} \frac{3}{N_1} \left(\sum_{k=0}^{\infty} (-1)^k \frac{(2k+1)!}{k!} c_{2k+1} \epsilon_{2k+1} \exp(-\epsilon_{2k+1} t) \right) \quad (B6)$$

with

$$N_1 = \sum_{k=0}^{\infty} (-1)^k \frac{(2k)!}{k!} c_{2k} \exp(-\epsilon_{2k} t) \quad (B7)$$

as normalization.

2. case $a_i < 0$ (maximum):

$$v^{(1)} = \frac{1}{N_2} \sum_{\substack{n=1 \\ n_1+\dots+n_d=n}}^{\infty} \epsilon_n e^{\epsilon_n t} \prod_{i=1}^d \left(\sqrt{\left| \frac{\Gamma(\frac{1}{4} + ib_i/2)}{\Gamma(\frac{3}{4} + ib_i/2)} \right|} \right) \sum_{k=1}^d I_k(\xi_k^2) \sqrt{\frac{D|a_k|}{2}} \quad (B8)$$

$$v_k^{(2)} = \frac{1}{N_2} \sum_{\substack{n=1 \\ n_1+\dots+n_d=n}}^{\infty} \epsilon_n e^{\epsilon_n t} \prod_{i=1}^d \left(\sqrt{\left| \frac{2\Gamma(\frac{3}{4} + ib_i/2)}{\Gamma(\frac{1}{4} + ib_i/2)} \right|} I_i(\cdot) \right) I_k(\xi_k) \sqrt[4]{\frac{|a_k|}{2D}} \quad (B9)$$

with

$$I_k(\xi_k) = \int_{-u_k}^{u_k} \xi_k y_1(\xi_k) d\xi_k \quad I_k(\xi_k^2) = \int_{-u_k}^{u_k} \xi_k^2 y_2(\xi_k) d\xi_k \quad I_k(\cdot) = \int_{-u_k}^{u_k} y_1(\xi_k) d\xi_k \quad (B10)$$

$$u_k = \frac{2U_m}{|a_k|} \sqrt[4]{\frac{|a_k|}{2D}} \quad b_k = \frac{n_k}{2\sqrt{D|a_k|}} \quad (B11)$$

$$N = \sum_{\substack{n=1 \\ n_1+\dots+n_d=n}}^{\infty} e^{\epsilon_n t} \prod_{i=1}^d \left(\sqrt{\left| \frac{\Gamma(\frac{1}{4} + ib_k/2)}{\Gamma(\frac{3}{4} + ib_k/2)} \right|} I_k(\cdot) \right) \quad (B12)$$

and functions $y_1()$, $y_2()$ defined in [19] page 692 as series

$$y_1(\xi_k) = 1 + b_k \frac{\xi_k^2}{2!} + \left(b_k^2 - \frac{1}{2}\right) \frac{\xi_k^4}{4!} + \dots \quad (\text{B13})$$

$$y_2(\xi_k) = \xi_k + b_k \frac{\xi_k^3}{3!} + \left(b_k^2 - \frac{3}{2}\right) \frac{\xi_k^5}{5!} + \dots \quad (\text{B14})$$

-
- [1] J.D. Nulton and P. Salamon. Statistical mechanics of combinatorial optimization. *Phys. Rev.*, **A 37**:1351, 1988.
- [2] B. Andresen. Finite-time thermodynamics and simulated annealing. In *Proceedings of the Fourth International Conference on Irreversible Processes and Selforganization*, Rostock, 1989.
- [3] P. Sibiani, K.M. Pedersen, K.H. Hoffmann, and P. Salamon. Monte Carlo dynamics of optimization: A scaling description. *Phys. Rev.*, **A 42**:7080, 1990.
- [4] H.P. Schwefel. *Evolution and optimum seeking*. Wiley, New York, 1995.
- [5] I. Rechenberg. *Evolutionsstrategien - Optimierung technischer Systeme nach Prinzipien der biologischen Information*. Friedrich Frommann Verlag (Günther Holzboog K.G.), Stuttgart - Bad Cannstatt, 1995.
- [6] R. Feistel and W. Ebeling. *Evolution of Complex Systems*. Kluwer Academic Publ., Dordrecht, 1989.
- [7] W. Ebeling and A. Engel. Models of Evolutionary systems and their applications to optimization problems. *Syst. Anal. Model. Simul.*, **3**:377, 1986.
- [8] T. Boseniuk, W. Ebeling, and A. Engel. Boltzmann and Darwin strategies in complex optimization. *Phys. Lett.*, **A 125**:307, 1987.
- [9] T. Boseniuk and W. Ebeling. Boltzmann-, Darwin- and Heackel-strategies in optimization problems. In *PPSN Dortmund*, 1990.
- [10] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. *J. Chem. Phys.*, **21**:1087, 1953.
- [11] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi. *Science*, **220**:671, 1983.
- [12] H. Risken. *The Fokker-Planck Equation*. Springer Verlag, 1989.
- [13] G. Jetschke. *Mathematik der Selbstorganisation*. Deutscher Verlag der Wissenschaften, 1989.
- [14] J. Milnor. *Morse theory*. Ann. of Mathematical Study **51**. Princeton Univ. Press, Princeton, 1963.
- [15] M. Conrad and W. Ebeling. M.V. Volkenstein, evolutionary thinking and the structue of fitness landscapes. *BioSystems*, **27**:125, 1992.
- [16] N. Berline, M. Vergne, and E. Getzler. *Heat kernels and Dirac Operators*. Springer Verlag, New York, 1992.
- [17] D. Husemoller. *Fiber Bundles*. MacGraw-Hill Book Co., New York-London-Sydney, 1966.
- [18] M. Atiyah. *K-Theory*. W.A. Benjamin, New York–Amsterdam, 1967.
- [19] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Applied Mathematical Series 55. National Bureau of Standards, December 1972.